

# SEQ SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2004, 10:00:01 ; Search time 3140 Seconds  
(without alignments)  
4582.767 Million cell updates/sec

Title: US-10-019-817C-2  
Perfect score: 1696  
Sequence: 1 MATIREVAKAAGVSPATVSR.....GTYKVTFP TKLVYRESCPKA 332

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US10019817/runat\_22072004\_091418\_11780/app\_query.fasta\_1.5  
19

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019817\_@CGN\_1\_1\_2527\_@runat\_22072004\_091418\_11780 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*

11: gb\_sts:\*  
 12: gb\_sy:\*  
 13: gb\_un:\*  
 14: gb\_vi:\*  
 15: em\_ba:\*  
 16: em\_fun:\*  
 17: em\_hum:\*  
 18: em\_in:\*  
 19: em\_mu:\*  
 20: em\_om:\*  
 21: em\_or:\*  
 22: em\_ov:\*  
 23: em\_pat:\*  
 24: em\_ph:\*  
 25: em\_pl:\*  
 26: em\_ro:\*  
 27: em\_sts:\*  
 28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		Query					
No.	Score	Match	Length	DB	ID	Description	
	1	1688	99.5	1250	1	AY040212 - GenBank 2001	AY040212 Lactobaci
	2	1688	99.5	1435	6	AX069168 wo 01/02576	AX069168 Sequence
	3	584	34.4	1005	6	AX570637	AX570637 Sequence
c	4	584	34.4	8136	6	AR218843	AR218843 Sequence
c	5	584	34.4	8136	6	BD003755	BD003755 Polynucle
	6	584	34.4	10393	1	AE008532	AE008532 Streptoco
	7	584	34.4	12841	1	AE007477	AE007477 Streptoco
	8	584	34.4	349980	6	AX571765	AX571765 Sequence
	9	579	34.1	1005	6	AR346815	AR346815 Sequence
	10	571	33.7	151947	2	SPNEU1902	AL449924 Streptoco
c	11	570	33.6	11926	1	AE014930	AE014930 Streptoco
c	12	568	33.5	5259	1	SMU21942	U21942 Streptococc
c	13	553.5	32.6	4199	1	STU61402	U61402 Streptococc
c	14	553.5	32.6	10954	1	AF389475	AF389475 Streptoco
c	15	553.5	32.6	14782	1	AF503446	AF503446 Streptoco
	16	550	32.4	6168	1	AF005933	AF005933 Lactobaci

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2004, 09:36:26 ; Search time 354 Seconds  
(without alignments)  
3984.186 Million cell updates/sec

Title: US-10-019-817C-2  
Perfect score: 1696  
Sequence: 1 MATIREVAKAAGVSPATVSR.....GTYKVTFTP TKLVYRESCPKA 332

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlh  
-

Q=/cgn2\_1/USPTO\_spool/US10019817/runat\_22072004\_091417\_11759/app\_query.fasta\_1.5  
19

-DB=N\_Geneseq\_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019817 @CGN\_1\_1\_352 @runat\_22072004\_091417\_11759 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	1688	99.5	1281	4	AAF56025	Aaf56025 Lactobaci
	2	1688	99.5	1435	4	AAF56001	Aaf56001 Lactobaci
	3	584	34.4	1005	7	ABX07636	Abx07636 S. pneumo
	4	584	34.4	1008	5	AAF29401	Aaf29401 DNA encod
c	5	584	34.4	8136	2	AAV52208	Aav52208 Streptoco
	6	584	34.4	110000	7	ABS56454_16	Continuation (17 o
	7	579	34.1	1005	9	ADC91799	Adc91799 E. faeciu
	8	511	30.1	1021	6	ABK74054	Abk74054 Bacillus
c	9	480.5	28.3	17764	2	AAX13238	Aax13238 Enterococ
c	10	480.5	28.3	17764	6	ABS99033	Abs99033 Enterococ
	11	425	25.1	1940	2	AAV37385	Aav37385 Streptoco
	12	416	24.5	3246	7	ABX14307	Abx14307 DNA encod
	13	409	24.1	4473	2	AAX13151	Aax13151 Enterococ
	14	409	24.1	4473	6	ABS98946	Abs98946 Enterococ
	15	378	22.3	1083	9	ADC90488	Adc90488 E. faeciu
	16	360.5	21.3	156638	6	ABQ81850	Abq81850 Bifidobac
c	17	339	20.0	1253	2	AAX30783	Aax30783 Streptoco
	18	319	18.8	1014	7	ACA19244	Aca19244 Prokaryot
	19	318	18.8	1107	5	AAH66350	Aah66350 C glutami
	20	318	18.8	1164	7	ACA00491	Aca00491 C. glutam
	21	318	18.8	1294	9	ADD13935	Add13935 C. glutam
	22	318	18.8	1600	6	AAL46365	Aal46365 C glutami
	23	318	18.8	349980	5	AAH68528	Aah68528 C glutami
	24	313.5	18.5	1005	7	ACA27829	Aca27829 Prokaryot
	25	313.5	18.5	1011	4	AAS53531	Aas53531 Haemophil
	26	313.5	18.5	1011	7	ACA34513	Aca34513 Prokaryot
	27	312	18.4	1011	7	ACA32028	Aca32028 Prokaryot
	28	308.5	18.2	1005	7	ACA42967	Aca42967 Prokaryot
	29	308	18.2	1023	7	ACA52154	Aca52154 Prokaryot
	30	304	17.9	1023	7	ACA50798	Aca50798 Prokaryot
	31	302.5	17.8	110000	2	AAT42063_16	Continuation (17 o
	32	301	17.7	990	7	ACA33441	Aca33441 Prokaryot
	33	301	17.7	1011	7	ACA35899	Aca35899 Prokaryot
	34	294	17.3	1011	6	ABN92795	Abn92795 Staphyloc
	35	289	17.0	1026	7	ACA51392	Aca51392 Prokaryot
	36	289	17.0	1029	7	ACA21642	Aca21642 Prokaryot
c	37	288.5	17.0	110000	6	ABA03041_16	Continuation (17 o
	38	284	16.7	1026	7	ACA19062	Aca19062 Prokaryot
	39	281.5	16.6	1038	5	AAH66475	Aah66475 C glutami
	40	281.5	16.6	1161	4	AAF71556	Aaf71556 Corynebac
	41	281.5	16.6	1500	6	AAL46362	Aal46362 C glutami
c	42	281.5	16.6	349980	5	AAH68528	Aah68528 C glutami
	43	279.5	16.5	1041	7	ACA27759	Aca27759 Prokaryot
c	44	279.5	16.5	110000	6	ABQ69245_16	Continuation (17 o
	45	279.5	16.5	110000	6	ABQ67195_0	Abq67195 Listeria

# ALIGNMENTS

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2004, 08:10:10 ; Search time 75 Seconds  
(without alignments)  
2456.583 Million cell updates/sec

Title: US-10-019-817C-2  
Perfect score: 1696  
Sequence: 1 MATIREVAKAAGVSPATVSR.....GTYKVTFPTKLVYRESCPKA 332

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlh  
-

Q=/cgn2\_1/USPTO\_spool/US10019817/runat\_22072004\_091415\_11693/app\_query.fasta\_1.5  
19

-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019817 @CGN\_1\_1\_54 @runat\_22072004\_091415\_11693 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
<hr/>							
c	1	584	34.4	8136	4	US-08-961-527-75	Sequence 75, Appl
	2	579	34.1	1005	4	US-09-107-532A-1426	Sequence 1426, Ap
	3	532.5	31.4	1011	4	US-09-134-000C-2831	Sequence 2831, Ap
	4	437.5	25.8	993	4	US-09-489-039A-2209	Sequence 2209, Ap
	5	378	22.3	1083	4	US-09-107-532A-115	Sequence 115, App
	6	365.5	21.6	1026	4	US-09-134-000C-2906	Sequence 2906, Ap
	7	314	18.5	1020	4	US-09-489-039A-5543	Sequence 5543, Ap
	8	307	18.1	1248	4	US-09-489-039A-2804	Sequence 2804, Ap
	9	302.5	17.8	1830121	4	US-09-557-884-1	Sequence 1, Appli
	10	302.5	17.8	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	11	294	17.3	1011	4	US-09-134-001C-2258	Sequence 2258, Ap
	12	294	17.3	1032	4	US-09-489-039A-2332	Sequence 2332, Ap
	13	287	16.9	1011	4	US-09-489-039A-4831	Sequence 4831, Ap
c	14	277	16.3	6846	4	US-08-961-527-198	Sequence 198, App
	15	274	16.2	1047	4	US-09-543-681A-693	Sequence 693, App
	16	270.5	15.9	996	4	US-09-489-039A-6861	Sequence 6861, Ap
	17	270.5	15.9	1170	4	US-09-489-039A-6057	Sequence 6057, Ap
	18	270	15.9	990	4	US-09-308-003-4	Sequence 4, Appli
	19	270	15.9	12173	4	US-08-956-171E-310	Sequence 310, App
	20	267	15.7	1131	4	US-09-543-681A-4105	Sequence 4105, Ap
	21	261	15.4	885	4	US-09-107-532A-1947	Sequence 1947, Ap
	22	252	14.9	4776	2	US-08-852-401-1	Sequence 1, Appli
	23	251.5	14.8	1026	4	US-09-543-681A-1803	Sequence 1803, Ap
	24	250	14.7	1125	4	US-09-489-039A-5027	Sequence 5027, Ap
	25	250	14.7	1197	4	US-09-428-082B-389	Sequence 389, App
	26	250	14.7	1221	4	US-09-025-769B-292	Sequence 292, App
	27	250	14.7	1341	4	US-09-522-666-3	Sequence 3, Appli
	28	250	14.7	3092	4	US-09-522-666-1	Sequence 1, Appli
	29	250	14.7	3150	4	US-09-522-666-5	Sequence 5, Appli
c	30	250	14.7	3604	3	US-09-235-246-20	Sequence 20, Appl
c	31	250	14.7	3740	1	US-08-041-648-3	Sequence 3, Appli
c	32	250	14.7	3740	1	US-08-217-529-3	Sequence 3, Appli
	33	250	14.7	3832	1	US-08-148-675A-2	Sequence 2, Appli
	34	250	14.7	3974	3	US-09-026-343-33	Sequence 33, Appl
	35	250	14.7	3974	3	US-09-042-105-16	Sequence 16, Appl
	36	250	14.7	3974	3	US-09-044-856A-7	Sequence 7, Appli
	37	250	14.7	3974	3	US-09-023-082A-147	Sequence 147, App
	38	250	14.7	3974	3	US-09-044-855A-7	Sequence 7, Appli
	39	250	14.7	3974	3	US-09-078-670-4	Sequence 4, Appli
	40	250	14.7	3974	4	US-09-026-408-14	Sequence 14, Appl
	41	250	14.7	3974	4	US-09-362-871-33	Sequence 33, Appl
	42	250	14.7	3974	4	US-09-627-154-4	Sequence 4, Appli
	43	250	14.7	3974	4	US-09-027-287-50	Sequence 50, Appl
	44	250	14.7	3974	4	US-09-437-602-4	Sequence 4, Appli
	45	250	14.7	3974	4	US-09-252-656B-50	Sequence 50, Appl

## ALIGNMENTS

RESULT 1  
US-08-961-527-75/c  
; Sequence 75, Application US/08961527

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2004, 08:56:01 ; Search time 398 Seconds  
(without alignments)  
4090.057 Million cell updates/sec

Title: US-10-019-817C-2  
Perfect score: 1696  
Sequence: 1 MATIREVAKAAGVSPATVSR.....GTYKVTFTPKLVYRESCPKA 332

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US10019817/runat\_22072004\_091417\_11726/app\_query.fasta\_1.5  
19

-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10019817@cgn\_1\_1\_354@runat\_22072004\_091417\_11726  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	% Query		DB	ID	Description
			Match	Length			
c	1	584	34.4	8136	13	US-10-158-844-75	Sequence 75, Appl
	2	511	30.1	1021	9	US-09-974-300-1345	Sequence 1345, Ap
c	3	480.5	28.3	17764	9	US-09-070-927A-301	Sequence 301, App
	4	409	24.1	4473	9	US-09-070-927A-214	Sequence 214, App
	5	360.5	21.3	2256646	17	US-10-470-565-1	Sequence 1, Appli
	6	338	19.9	1062	15	US-10-156-761-5304	Sequence 5304, Ap
c	7	338	19.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
	8	319	18.8	1014	13	US-10-282-122A-7114	Sequence 7114, Ap
	9	318	18.8	1107	9	US-09-738-626-1385	Sequence 1385, Ap
	10	318	18.8	1600	9	US-09-938-540-1	Sequence 1, Appli
	11	318	18.8	3309400	9	US-09-738-626-1	Sequence 1, Appli
	12	313.5	18.5	1005	13	US-10-282-122A-15699	Sequence 15699, A
	13	313.5	18.5	1011	9	US-09-815-242-7168	Sequence 7168, Ap
	14	313.5	18.5	1011	13	US-10-282-122A-22383	Sequence 22383, A
	15	312	18.4	1011	13	US-10-282-122A-19898	Sequence 19898, A
	16	310.5	18.3	1044	15	US-10-156-761-819	Sequence 819, App
	17	308.5	18.2	1005	13	US-10-282-122A-30837	Sequence 30837, A
	18	308	18.2	1023	13	US-10-282-122A-40024	Sequence 40024, A
	19	304	17.9	1023	13	US-10-282-122A-38668	Sequence 38668, A
	20	302.5	17.8	1830121	15	US-10-329-960-1	Sequence 1, Appli
	21	302.5	17.8	1830121	16	US-10-329-670-1	Sequence 1, Appli
	22	301	17.7	990	13	US-10-282-122A-21311	Sequence 21311, A
	23	301	17.7	1011	13	US-10-282-122A-23769	Sequence 23769, A
	24	295	17.4	1020	15	US-10-156-761-547	Sequence 547, App
	25	295	17.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
	26	289	17.0	1026	13	US-10-282-122A-39262	Sequence 39262, A
	27	289	17.0	1029	13	US-10-282-122A-9512	Sequence 9512, Ap
	28	288.5	17.0	1047	15	US-10-156-761-6265	Sequence 6265, Ap
	29	284	16.7	1026	13	US-10-282-122A-6932	Sequence 6932, Ap
	30	281.5	16.6	1038	9	US-09-738-626-1510	Sequence 1510, Ap
	31	281.5	16.6	1500	9	US-09-938-642-1	Sequence 1, Appli
	32	281.5	16.6	3309400	9	US-09-738-626-1	Sequence 1, Appli
c	33	279.5	16.5	1041	13	US-10-282-122A-15629	Sequence 15629, A
	34	279.5	16.5	495269	16	US-10-398-221-8	Sequence 8, Appli
c	35	279.5	16.5	3011208	16	US-10-398-221-2058	Sequence 2058, Ap
	36	277	16.3	1113	17	US-10-474-776-565	Sequence 565, App
c	37	277	16.3	6846	13	US-10-158-844-198	Sequence 198, App
	38	276.5	16.3	1035	13	US-10-282-122A-41112	Sequence 41112, A
	39	275.5	16.2	1044	15	US-10-156-761-1684	Sequence 1684, Ap



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 28, 2004, 10:43:02 ; Search time 2366 Seconds  
(without alignments)  
4190.298 Million cell updates/sec

Title: US-10-019-817C-2  
Perfect score: 1696  
Sequence: 1 MATIREVAKAAGVSPATVSR.....GTYKVTFTP TKLVYRESCPKA 332

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delect 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

-  
Q=/cgn2\_1/USPTO\_spool/US10019817/runat\_22072004\_091418\_11819/app\_query.fasta\_1.5  
19

-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019817 @CGN 1 1 2607 @runat\_22072004\_091418\_11819 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*

12: gb\_est3:\*  
 13: gb\_est4:\*  
 14: gb\_est5:\*  
 15: em\_estfun:\*  
 16: em\_estom:\*  
 17: em\_gss\_hum:\*  
 18: em\_gss\_inv:\*  
 19: em\_gss\_pln:\*  
 20: em\_gss\_vrt:\*  
 21: em\_gss\_fun:\*  
 22: em\_gss\_mam:\*  
 23: em\_gss\_mus:\*  
 24: em\_gss\_pro:\*  
 25: em\_gss\_rod:\*  
 26: em\_gss\_phg:\*  
 27: em\_gss\_vrl:\*  
 28: gb\_gss1:\*  
 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		%						
Result			Query					
No.	Score	Match	Length	DB	ID			Description
c	1	204.5	12.1	565	28	BZ336973		BZ336973 ia84d06.g
	2	200	11.8	710	28	AY080106		AY080106 AY080106
c	3	193	11.4	579	9	AL043868		AL043868 DKFZp434C
c	4	190	11.2	589	9	AL037742		AL037742 DKFZp564C
	5	189	11.1	1081	28	AF095417		AF095417 AF095417
c	6	187.5	11.1	498	9	AL039076		AL039076 DKFZp566G
c	7	180.5	10.6	569	9	AL039077		AL039077 DKFZp566G
c	8	180.5	10.6	723	9	AL039416		AL039416 DKFZp434L
c	9	176	10.4	975	28	AF075804		AF075804 AF075804
c	10	175.5	10.3	521	14	CA890032		CA890032 B0156F05-
c	11	174.5	10.3	528	9	AL038811		AL038811 DKFZp566O
	12	170.5	10.1	750	29	CG898460		CG898460 pastbac07
c	13	168.5	9.9	740	9	AL042909		AL042909 DKFZp434J
c	14	168.5	9.9	794	9	AL045353		AL045353 DKFZp434B
c	15	167.5	9.9	378	14	CF307719		CF307719 ABF--01-E
c	16	167.5	9.9	458	14	CA887583		CA887583 B0137H04-
c	17	167.5	9.9	577	14	CA886289		CA886289 B0125H09-
c	18	167.5	9.9	720	14	CA881974		CA881974 K0994C10-
	19	163.5	9.6	584	9	AV609956		AV609956 AV609956
c	20	163.5	9.6	734	9	AL048427		AL048427 DKFZp586H
	21	159	9.4	679	28	BZ561058		BZ561058 pacs2-164
c	22	154.5	9.1	1025	9	AL038025		AL038025 DKFZp566C
c	23	154	9.1	753	9	AL045341		AL045341 DKFZp434B
c	24	153.5	9.1	621	14	CA895570		CA895570 B0193A03-
	25	153.5	9.1	682	10	AW200635		AW200635 da23h03.y
c	26	152.5	9.0	774	9	AL045337		AL045337 DKFZp434A
c	27	149	8.8	447	28	BH400232		BH400232 AG-ND-153
c	28	148.5	8.8	362	14	CF337062		CF337062 JMT--07-F
	29	147.5	8.7	1519	28	BZ568229		BZ568229 pacs2-164

# SEO SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 06:55:55 ; Search time 2062 Seconds  
(without alignments)  
2059.950 Million cell updates/sec

Title: US-10-019-817C-9  
Perfect score: 98  
Sequence: 1 tgtttactaaaaatattttg.....aacacctaaaggagaaaatc 98

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%	Query						
No.	Score	Match	Length	DB	ID			Description	
1	98	100.0	98	6	AX069176	wo 01/02576		AX069176 Sequence	
2	98	100.0	98	6	AX069189	"		AX069189 Sequence	
3	98	100.0	962	1	AY040211	GenBank 2001		AY040211 Lactobaci	
4	74.2	75.7	1102	6	I46883			I46883 Sequence 2	
5	74.2	75.7	5015	1	LBALACZBUL			M55068 L.delbrueck	
6	39.6	40.4	160867	9	AC025254			AC025254 Homo sapi	
c 7	39.6	40.4	205328	2	AC024467			AC024467 Homo sapi	
8	36.4	37.1	150572	9	CNS01RIC			AL163151 Human chr	
9	33.6	34.3	159618	9	AC096750			AC096750 Homo sapi	
c 10	33.2	33.9	79572	8	AP005719			AP005719 Oryza sat	
11	33.2	33.9	157616	2	AC123599			AC123599 Mus muscu	
12	33.2	33.9	202290	10	AC127561			AC127561 Mus muscu	
c 13	33	33.7	181062	10	AC135114			AC135114 Mus muscu	
c 14	33	33.7	224941	10	AC122933			AC122933 Mus muscu	
15	32.2	32.9	1214	3	AK116288			AK116288 Ciona int	
16	32.2	32.9	154918	9	HSDJ67A8			AL121957 Human DNA	
17	32.2	32.9	155362	2	AC022244			AC022244 Homo sapi	
c 18	32	32.7	4152	9	HSM806683			BX640637 Homo sapi	
19	32	32.7	110909	9	AC117510			AC117510 Homo sapi	
c 20	32	32.7	120178	9	AC114760			AC114760 Homo sapi	
21	32	32.7	188644	9	AC090518			AC090518 Homo sapi	
c 22	32	32.7	207223	10	AL732521			AL732521 Mouse DNA	
c 23	32	32.7	211901	10	AL772310			AL772310 Mouse DNA	
c 24	32	32.7	241961	2	AC111911			AC111911 Rattus no	
c 25	31.8	32.4	89453	8	AC068655			AC068655 Genomic S	
c 26	31.8	32.4	107579	8	AC016162			AC016162 Arabidops	
c 27	31.8	32.4	155641	2	AC125105			AC125105 Mus muscu	
c 28	31.8	32.4	251952	10	AL591865			AL591865 Mouse DNA	
29	31.6	32.2	2418	3	DDRASGG			Z11533 D.discoideu	
c 30	31.6	32.2	83373	8	AB017064			AB017064 Arabidops	
31	31.6	32.2	97579	2	AC146722			AC146722 Medicago	
c 32	31.6	32.2	177994	2	AC119888			AC119888 Mus muscu	
c 33	31.6	32.2	186250	2	AC132271			AC132271 Mus muscu	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 06:55:05 ; Search time 305 Seconds  
(without alignments)  
1364.995 Million cell updates/sec

Title: US-10-019-817C-9  
Perfect score: 98  
Sequence: 1 tgtttactaaaaatattttg.....aacacctaaaggagaaaatc 98

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	98	100.0	98	4	AAF56005	Aaf56005 Lactobaci
2	98	100.0	98	4	AAF56018	Aaf56018 Lactobaci
3	98	100.0	107	4	AAF56023	Aaf56023 Lactobaci
4	74.2	75.7	106	4	AAF56024	Aaf56024 Lactobaci
5	74.2	75.7	1102	2	AAQ88812	Aaq88812 Clone pDP
6	57	58.2	57	4	AAF56019	Aaf56019 Lac promo
7	50.6	51.6	57	4	AAF56020	Aaf56020 Lac promo
8	45.8	46.7	57	4	AAF56021	Aaf56021 Lactobaci

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 00:23:02 ; Search time 39 Seconds  
(without alignments)  
1394.492 Million cell updates/sec

Title: US-10-019-817C-9  
Perfect score: 98  
Sequence: 1 tggtttactaaaaatatatttg.....aacacctaagagagaaaatc 98

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%				ID	Description
	No.	Score	Match	Length	DB	Query		
c	1	74.2	75.7	1102	1	US-08-297-294A-2	Sequence 2, Appli	
	2	30.2	30.8	998	3	US-09-122-400B-5	Sequence 5, Appli	
	3	29	29.6	96109	4	US-09-596-002-35	Sequence 35, Appl	
	4	28.4	29.0	2635	3	US-08-714-918-38	Sequence 38, Appl	
	5	28.4	29.0	2635	3	US-09-265-315-38	Sequence 38, Appl	
	6	28.4	29.0	2635	3	US-09-265-315-38	Sequence 38, Appl	
	7	28.4	29.0	2635	3	US-09-266-417-38	Sequence 38, Appl	
	8	28.4	29.0	2635	4	US-09-528-709-38	Sequence 38, Appl	
	9	28.4	29.0	2635	4	US-09-527-745-38	Sequence 38, Appl	
	10	28.4	29.0	786431	4	US-09-751-389-3	Sequence 3, Appli	
	11	28.2	28.8	1866	4	US-09-601-198-153	Sequence 153, App	
	12	28.2	28.8	640681	4	US-09-790-988-1	Sequence 1, Appli	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 06:22:10 ; Search time 319 Seconds  
(without alignments)  
1506.294 Million cell updates/sec

Title: US-10-019-817C-9  
Perfect score: 98  
Sequence: 1 tgtttactaaaaatattttg.....aacacctaaggagaaaatc 98

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	% Query						Description
	No.	Score	Match	Length	DB	ID	
c	1	30.2	30.8	998	9	US-09-816-894-5	Sequence 5, Appli
c	2	30	30.6	544	17	US-10-021-323-10126	Sequence 10126, A
c	3	29.8	30.4	344	13	US-10-424-599-136179	Sequence 136179,
	4	29.6	30.2	3673778	15	US-10-312-841-2	Sequence 2, Appli
	5	29.2	29.8	728	13	US-10-027-632-24519	Sequence 24519, A
	6	29.2	29.8	728	13	US-10-027-632-24520	Sequence 24520, A
	7	29.2	29.8	728	16	US-10-027-632-24519	Sequence 24519, A
	8	29.2	29.8	728	16	US-10-027-632-24520	Sequence 24520, A
	9	29.2	29.8	1522	13	US-10-282-122A-10588	Sequence 10588, A
c	10	29.2	29.8	80332	13	US-10-087-192-1492	Sequence 1492, Ap
	11	29	29.6	952	13	US-10-424-599-130981	Sequence 130981,
	12	29	29.6	96109	13	US-10-672-787-35	Sequence 35, Appl
c	13	28.8	29.4	346	17	US-10-437-963-45580	Sequence 45580, A
c	14	28.8	29.4	3157	17	US-10-437-963-58235	Sequence 58235, A
	15	28.8	29.4	8033	13	US-10-221-613-73	Sequence 73, Appl
	16	28.8	29.4	8033	13	US-10-221-714A-29	Sequence 29, Appl
c	17	28.8	29.4	59463	16	US-10-085-117-184	Sequence 184, App
	18	28.8	29.4	312477	17	US-10-317-883A-12	Sequence 12, Appl
	19	28.6	29.2	418	13	US-10-085-783A-38657	Sequence 38657, A
	20	28.6	29.2	418	16	US-10-242-535A-38657	Sequence 38657, A
c	21	28.6	29.2	713	13	US-10-027-632-8804	Sequence 8804, Ap
c	22	28.6	29.2	713	16	US-10-027-632-8804	Sequence 8804, Ap
c	23	28.6	29.2	109586	13	US-10-087-192-220	Sequence 220, App
	24	28.4	29.0	516	17	US-10-437-963-29659	Sequence 29659, A
	25	28.4	29.0	942	9	US-09-815-242-4561	Sequence 4561, Ap
	26	28.4	29.0	945	9	US-09-815-242-8433	Sequence 8433, Ap
	27	28.4	29.0	945	9	US-09-815-242-8779	Sequence 8779, Ap
	28	28.4	29.0	1561	9	US-09-729-674-177	Sequence 177, App
	29	28.4	29.0	1686	16	US-10-369-493-26272	Sequence 26272, A
	30	28.4	29.0	6494	15	US-10-311-455-1365	Sequence 1365, Ap
	31	28.4	29.0	73334	15	US-10-311-455-2097	Sequence 2097, Ap
	32	28.4	29.0	73334	17	US-10-240-589C-127	Sequence 127, App
	33	28.4	29.0	786431	15	US-10-412-277-3	Sequence 3, Appli
c	34	28.4	29.0	3673778	15	US-10-312-841-2	Sequence 2, Appli
c	35	28.2	28.8	396	12	US-09-732-627A-781	Sequence 781, App
c	36	28.2	28.8	476	15	US-10-198-846-12033	Sequence 12033, A
	37	28.2	28.8	626	15	US-10-157-031-412	Sequence 412, App
	38	28.2	28.8	656	13	US-10-027-632-206777	Sequence 206777,
	39	28.2	28.8	656	13	US-10-027-632-206778	Sequence 206778,
	40	28.2	28.8	656	13	US-10-027-632-206779	Sequence 206779,
	41	28.2	28.8	656	16	US-10-027-632-206777	Sequence 206777,
	42	28.2	28.8	656	16	US-10-027-632-206778	Sequence 206778,
	43	28.2	28.8	656	16	US-10-027-632-206779	Sequence 206779,
c	44	28.2	28.8	680	15	US-10-198-846-3965	Sequence 3965, Ap
c	45	28.2	28.8	712	13	US-10-027-632-148502	Sequence 148502,

## ALIGNMENTS

RESULT 1

US-09-816-894-5/c



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 07:49:10 ; Search time 2403 Seconds  
(without alignments)  
1217.850 Million cell updates/sec

Title: US-10-019-817C-9  
Perfect score: 98  
Sequence: 1 tgtttactaaaaatatatttg.....aacacctaaaggagaaaatc 98

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gssl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%			Query			DB	ID	Description
	No.	Score	Match	Length	DB	ID			
c	1	36.4	37.1	266	28	AQ079030			AQ079030 CIT-HSP-2
	2	33.6	34.3	895	28	AZ527570			AZ527570 ENTCU11TF
	3	33.2	33.9	483	28	BH463398			BH463398 BOHFS68TR
c	4	33	33.7	456	28	AZ294388			AZ294388 RPCI-23-1
c	5	32.8	33.5	934	9	AL664954			AL664954 AL664954
c	6	32.6	33.3	518	28	AQ188655			AQ188655 HS_3231_B
c	7	32.4	33.1	544	10	BB730826			BB730826 BB730826
	8	32.2	32.9	243	13	BW200859			BW200859 BW200859
	9	32.2	32.9	426	28	AQ237925			AQ237925 RPCI11-64
c	10	32.2	32.9	529	10	AW702828			AW702828 TgESTzz89
c	11	32.2	32.9	536	12	BM133005			BM133005 TgESTzyb1
	12	32.2	32.9	577	13	BW040098			BW040098 BW040098
	13	32.2	32.9	674	28	CC138948			CC138948 NDL.4N2.S
	14	32	32.7	368	12	BM351953			BM351953 ig84f03.x
	15	32	32.7	707	28	BH948064			BH948064 obu93g12.
c	16	31.8	32.4	560	14	CD766490			CD766490 AGENCOURT
c	17	31.8	32.4	900	28	CC074563			CC074563 CSU-K33r.
c	18	31.6	32.2	480	9	AA872410			AA872410 oh74e08.s
	19	31.6	32.2	581	13	BW165301			BW165301 BW165301
c	20	31.6	32.2	958	14	CD301051			CD301051 AGENCOURT
c	21	31.4	32.0	879	28	BZ976978			BZ976978 PUCAR07TD
	22	31.4	32.0	1160	13	BX440026			BX440026 BX440026
	23	31.2	31.8	343	13	BU795311			BU795311 SJF2DNB03
	24	31.2	31.8	533	10	BB696376			BB696376 BB696376
	25	31.2	31.8	663	29	AG164091			AG164091 Pan trogl
c	26	31.2	31.8	724	13	BU723823			BU723823 SJMBAE12
	27	31	31.6	224	9	AA562295			AA562295 vk97a12.r
	28	31	31.6	472	28	B25692			B25692 F28N22TR IG
	29	31	31.6	698	28	BH947699			BH947699 obv10f10.
c	30	31	31.6	752	28	BH731649			BH731649 BOMES70TR
	31	31	31.6	783	13	BW147202			BW147202 BW147202
	32	31	31.6	1201	9	AL559918			AL559918 AL559918
	33	30.8	31.4	294	28	CC091859			CC091859 CSU-K34.1
	34	30.8	31.4	449	12	BI384375			BI384375 BFLG2_002
c	35	30.8	31.4	547	9	AA604223			AA604223 no87a03.s
c	36	30.6	31.2	288	9	AA505808			AA505808 nh98e04.s
	37	30.6	31.2	363	9	AA821774			AA821774 vp77e07.r
	38	30.6	31.2	363	29	CG054009			CG054009 PUJAW67TD
	39	30.6	31.2	489	28	AQ219615			AQ219615 HS_3253_A
c	40	30.6	31.2	498	28	B89142			B89142 CIT-HSP-217
	41	30.6	31.2	526	28	AQ763179			AQ763179 HS_3222_B
	42	30.6	31.2	620	10	AW689139			AW689139 NF015G11S
	43	30.6	31.2	672	11	AY066481			AY066481 Schmidtea
	44	30.6	31.2	692	13	BX859859			BX859859 BX859859
	45	30.6	31.2	716	10	AW584488			AW584488 N210538e